

GERALD TAN

Mobile: +65 96673275 Email: tanjunweigerald1997@gmail.com

Github: <https://github.com/Geralddtan>

LinkedIn: <https://www.linkedin.com/in/gerald-tan/>

EDUCATION

National University of Singapore (NUS)

Aug 18 – present

- Bachelor of Science in Business Analytics (School of Computing), Honours | Minor in Statistics | CAP: 4.70/5.00
- Recipient of Bond Free NUS Merit Scholarship for outstanding academic excellence and leadership qualities
- Expected Date of Graduation: May 2022

NUS School of Computing, Research Assistant (Social Media Network Analysis)

Jan 21 – May 21

- Performed data collection, feature engineering using Python and graph network visualisation & analysis using Gephi to craft 3 datasets for undergraduate students taking the class

NUS School of Computing, Teaching Assistant (Introduction to Business Analytics)

Aug 19 – Dec 20

- Coached weekly sessions to clarify students' doubts on fundamentals of Business Analytics

TECHNICAL SKILLS

- Proficient in Python, Pyspark, HDFS, R, Java, SQL and cloud computing using GCP
- Proficient in software visualisation tools with Tableau, QGIS
- Proficient in Data Science Models, Data Management and Data Visualisation

WORK EXPERIENCE

Shopee, Data Science Intern

Jan 21 – Jun 21

- Engineered an airflow pipeline for Pyspark scripts to automate monthly manual reruns saving about 15 hours of work each month
- Performed geospatial analysis and visualisation with QGIS and KeplerGL to derive insights on user locations over 3 countries – Vietnam, Indonesia, Malaysia
- Coordinated with multiple stakeholders from various backgrounds to engineer an ETL Pyspark pipeline – complex feature engineering, merging 7 sources of data before ingestion into Elastic Search for live deployment
- Constructed complex elastic search matching queries for optimal query searches, achieving 96% accuracy after benchmarking - similar to Google's geocoding API
- Employed Locust as API load test framework on geocoding & autocomplete API before live deployment

NUS Fintech Society, Technology Analyst, Machine Learning

Aug 20 – May 21

- Built a Natural Language Processing (NLP) Web application scraping top news data from top 30 Nasdaq companies, calculating its sentiment scores based on recent news
- Employed React & Django with pre-trained Google Finbert, a pre-trained model designed to analyse sentiment of financial text

Lynx Analytics, Data Science Intern

May 20 – Aug 20

- Developed an automated Java & Python pipeline for large graph visualisation using Gephi Java Toolkit and DataShader
- Authored 2 analytical articles for marketing outreach on cryptocurrency hacks ([Upbit Hack](#)) and on growth of [Stable Coins](#) requiring business thinking, analytical insights and customer centric visualisations
- Pitched and garnered keen interest from 2 new clients globally (Monetary Authority of Singapore & Nexo)
- Utilised cloud technologies (GCP) on blockchain cryptocurrency transaction data to increase computation speed

ASTAR Genome Institute of Singapore, Data Science Intern

May 19 – Jul 19

- Engineered an Expectation-Maximisation algorithm package to iteratively infer microbiome interactions from microbiome abundance and external perturbation data
- Designed and developed a simulation pipeline for model testing, attained an accuracy of up to 97%
- Awarded Joint Authorship for Manuscript accepted into PLOS Computational Biology - a prestigious journal

COMPETITIONS & PROJECTS

Winner - LifeHack2020 NUS School of Computing Hackathon in collaboration with Shopee, Holmusk

- Worked in a team of 2, achieving first place in the Singapore wide healthcare analytics and visualisation hackathon focussing on a 3 step method to aid patients and doctors in breast cancer detection and awareness
- Trained an ensemble model using logistic regression and Resnet18 CNN model for breast cancer detection
- Designed a Tableau dashboard to raise awareness for breast cancer in Singapore context

1st Runner Up – NTU iNTUition v7.0 Hackathon

- Worked in a team of 4, attaining 1st runner up at iNTUition v7.0 Hackathon, one of Southeast Asia's largest student run hackathons
- Developed an automated cheat sheet builder webpage with React, ExpressJS and Flask - utilising OpenCV East CNN model for text font size detection and a custom 2D packing algorithm for optimal packing as final output

Automated Google Calendar appointments via Telegram Bot

- Engineered a telegram bot with Natural Language Processing abilities to automatically create google calendar events based on user event input, via Google API

Analytics Vidhya Hackathon (<http://bit.ly/37baQdR>)

- Competed in an Online Analytics Vidhya Hackathon for Sales Data with Python and Tableau
- Attained a ranking of 143 out of 14000 global registered participants, a field of more experienced data analysts
- Implemented Lasso Polynomial Regression algorithm using GridSearchCV for parameter optimisation

CO-CURRICULAR ACTIVITIES

Tembusu College, Dodgeball Exco Committee

Feb 19 – May 20

- Organised training sessions for more than 30 members, leading team to 2nd place in the 2019 Inter College Games

Organising Committee, Community Development Leadership Project

Jan 18 – May 20

- Lead and organised a community kitchen in Punggol to engage elderly residents living in isolation through organised group activities and interaction